A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP9009.

Evaluation of a Gene–Environment Interaction of *PON1* and Low-Level Nerve Agent Exposure with Gulf War Illness: A Prevalence Case–Control Study Drawn from the U.S. Military Health Survey's National Population Sample

Robert W. Haley,[1] *Gerald Kramer,*[1] *Junhui Xiao,*[1] *Jill A. Dever,*[2] *and John F. Teiber*[1]

[1]Division of Epidemiology, Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas, USA
[2]RTI International, Washington, District of Columbia, USA

Background: Consensus on the etiology of 1991 Gulf War illness (GWI) has been limited by lack of objective individual-level environmental exposure information and assumed recall bias.

Objectives: We investigated a prestated hypothesis of the association of GWI with a gene–environment (GxE) interaction of the *paraoxonase-1* (*PON1*) Q192R polymorphism and low-level nerve agent exposure.

Methods: A prevalence sample of 508 GWI cases and 508 nonpaired controls was drawn from the 8,020 participants in the U.S. Military Health Survey, a representative sample survey of military veterans who served during the Gulf War. The *PON1* Q192R genotype was measured by real-time polymerase chain reaction (RT-PCR), and the serum Q and R isoenzyme activity levels were measured with PON1-specific substrates. Low-level nerve agent exposure was estimated by survey questions on having heard nerve agent alarms during deployment.

Results: The GxE interaction of the Q192R genotype and hearing alarms was strongly associated with GWI on both the multiplicative [prevalence odds ratio (POR) of the interaction = 3.41; 95% confidence interval (CI): 1.20, 9.72] and additive (synergy index = 4.71; 95% CI: 1.82, 12.19) scales, adjusted for measured confounders. The Q192R genotype and the alarms variable were independent (adjusted POR in the controls = 1.18; 95% CI: 0.81, 1.73; $p = 0.35$), and the associations of GWI with the number of R alleles and quartiles of Q isoenzyme were monotonic. The adjusted relative excess risk due to interaction (aRERI) was 7.69 (95% CI: 2.71, 19.13). Substituting Q isoenzyme activity for the genotype in the analyses corroborated the findings. Sensitivity analyses suggested that recall bias had forced the estimate of the GxE interaction toward the null and that unmeasured confounding is unlikely to account for the findings. We found a GxE interaction involving the Q-correlated PON1 diazoxonase activity and a weak possible GxE involving the Khamisiyah plume model, but none involving the PON1 R isoenzyme activity, arylesterase activity, paraoxonase activity, butyrylcholinesterase genotypes or enzyme activity, or pyridostigmine.

Discussion: Given gene–environment independence and monotonicity, the unconfounded aRERI > 0 supports a mechanistic interaction. Together with the direct evidence of exposure to fallout from bombing of chemical weapon storage facilities and the extensive toxicologic evidence of biochemical protection from organophosphates by the Q isoenzyme, the findings provide strong evidence for an etiologic role of low-level nerve agent in GWI. https://doi.org/10.1289/EHP9009

Introduction

In the 1991 Persian Gulf War, approximately 700,000 U.S. military personnel and 300,000 people from 41 Coalition countries were deployed to the Kuwaiti Theater of Operations (KTO) for a 5-wk air war punctuated by a 5-d ground war.[1] For months after the short deployment, tens of thousands of previously fit personnel developed an often-disabling set of symptoms, termed Gulf War illness (GWI), including fatigue, memory and concentration impairment, difficulty finding words, insomnia, diarrhea or constipation, cutaneous tingling and numbness, balance disturbance and vertigo attacks, body temperature dysregulation, and often severe somatic pain,[2–4] which have persisted.[5] Rates of these symptoms were higher in the KTO-deployed than in the nondeployed U.S. force.[6,7] Among the deployed, both combat and support personnel were affected,[8–10] and psychological explanations do not fully explain the illness.[11] Clinical case–control studies employing neuroimaging, electroencephalography, and autonomic testing have identified abnormalities of brain and peripheral nerve function or metabolism underlying the symptoms.[12–20]

In the first published epidemiological study of environmental risk factors completed 4 y after the war ($n = 249$), our group found the strongest associations of GWI with self-reported low-level organophosphate nerve agent exposure and having experienced adverse effects of antinerve gas tablets containing the carbamate pyridostigmine bromide.[21] To try to explain why only a fraction of those exposed to these agents developed GWI, we followed up with a prevalence case–control study ($n = 40$) drawn from the first cohort, in which we found GWI inversely associated with serum activity of the Q isoenzyme of the *paraoxonase-1 (PON1)* gene, a known genetic determinant of susceptibility to organophosphate cholinesterase-inhibiting chemicals including nerve agents.[22]

The PON1 enzyme hydrolyzes several important substrate molecules such as paraoxon (the active metabolite of the pesticide parathion) and diazoxon (the active metabolite of the pesticide diazinon) as well as nerve agents like sarin and soman. A given subject's PON1 enzyme hydrolyzes these substrates at very different levels of catalytic efficiency. For example, one's PON1 enzyme can have very high catalytic activity against sarin (high sarinase activity) but very low activity against paraoxon (low paraoxonase activity). The enzyme was named "paraoxonase" after the first substrate it was found to hydrolyze.

The *PON1* gene contains a common polymorphism in codon 192 that directs the production of either the 192 glutamine (Q) isoenzyme or the 192 arginine (R) isoenzyme, the only catalytic enzymes in humans that hydrolyze, and thus inactivate, organophosphates. QQ homozygous individuals produce only the Q isoenzyme, which efficiently hydrolyzes nerve agents like sarin; RR homozygotes produce only the R isoenzyme, which is relatively ineffective against nerve agents; and QR heterozygotes produce

Address correspondence to Robert W. Haley, Department of Internal Medicine, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, Texas 75390 USA. Email: Robert.Haley@UTSouthwestern.edu

Supplemental Material is available online (https://doi.org/10.1289/EHP9009).

The authors declare they have no actual or potential competing financial interests.

Received 21 January 2021; Revised 23 February 2022; Accepted 21 March 2022; Published 11 May 2022.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

variable proportions of both[23,24]; moreover, within genotype the level of Q isoenzyme activity varies >10-fold.[23,25] As hypothesized, we found GWI was significantly elevated in veterans with an R allele (RR or QR genotypes) and in those with lower serum activity of the Q isoenzyme—a pattern compatible with an increased susceptibility to nerve agents.[22] The sample size was too small to test for a gene–environment (GxE) interaction.

The butyrylcholinesterase (BChE) enzyme (serum cholinesterase, pseudocholinesterase) normally contributes to protection from organophosphates and pyridostigmine by covalently binding and sequestering them. An early case report suggested that variants of the *BChE* gene with lower organophosphate binding activity may have contributed to GWI.[26] Our study[22] and a later one[27] found no association of GWI with genetic variants or the serum activity level of BChE, but the Steele et al. study[27] suggested that the uncommon *BChE* gene variants K/K, U/AK, U/A, A/F, and AK/F may have modified the association of GWI with having taken pyridostigmine antinerve agent tablets.

Despite subsequent evidence further linking GWI with widespread exposure to cholinesterase-inhibiting chemicals, reviewed exhaustively by Michalovicz et al.,[28] no consensus on the role of these environmental exposures has developed because of common study design flaws, including small, unrepresentative samples of veterans; self-selection of volunteer participants; and post hoc exploratory analyses of multiple risk factors.[29] The criticism most often cited has been the assumption that recall bias from self-reported environmental exposure measures inflated their associations with GWI.[30]

We thus undertook the present study to test the prestated hypothesis that, if low-level organophosphate nerve agent exposure caused GWI, nerve agent–exposed veterans homozygous for the minor RR genotype (having no Q isoenzyme) and those QQ or QR individuals with lower levels of Q isoenzyme activity would have higher rates of GWI. To overcome the challenges, we performed a large population-representative random sample survey, measured *PON1* and *BChE* genotypes and enzyme activity levels in a large prevalence case–control subsample drawn from the survey participants, tested the primary prestated hypothesis of a GxE interaction between the *PON1* Q192R genotype and low-level nerve agent exposure on both the additive and multiplicative scales, controlled selection bias by random sample selection and confounding by multivariable analysis, and addressed recall bias and unmeasured confounding by sensitivity testing.

Methods

National Survey of Gulf War-Era Veterans

To obtain information from a representative sample of veterans, from 2007 to 2010 we conducted a national prevalence survey known as the U.S. Military Health Survey (USMHS) via a computer-assisted telephone interview (CATI) with 8,020 veterans.[7] The original sample of 14,812 veterans was randomly selected from the target population of 3,492,407 on active duty, National Guard, and Reserves in the personnel file of the Gulf War-era military population covering 2 August 1990 to 1 July 1991 maintained by the Defense Manpower Data Center, Seaside, California. We stratified the personnel file by the following design parameters prior to sample selection: a flag indicating deployment to the KTO; age (<49 y, ≥49 y); sex; race/ethnicity (non-Hispanic White vs. other to ensure adequate representation of minority groups in the sample); military rank during the war (officer, enlisted); military component (active duty, Reserve/Guard); unit location in KTO on 20 January 1991 (deployed only) and special study strata, including twin pairs; members of the 24th Reserve Naval Mobile Construction Battalion; and parents of a child with Goldenhar complex birth defect (Figure 1). With 74.9% of the randomly selected veterans located and contacted and 80.2% of these agreeing to participate, the overall response rate was 60.1%.[31] Of the full USMHS sample, 6,497 were deployed to the KTO, and 1,523 were nondeployed. The survey methods, extensive pilot testing and initial findings of the USMHS were described in detail elsewhere.[7]

Case Definition

The original GWI Research case definition, used in this study, was developed from a survey of symptoms collected within 4 y of the end of the Gulf War in deployed members of a Naval Reserve construction battalion.[2] To reduce ambiguity in symptom reporting, for each of the 27 typical GWI symptoms the questionnaire included a battery of 4–11 clarifying questions that were analyzed by principal components factor analysis to derive 2–3 unambiguous symptom scales. The resulting 52 continuous symptom component scales were analyzed with a second-stage factor analysis that identified 6 strong latent factors, suggesting possible GWI variants. The resulting six factor scales were dichotomized at 1.5 standard deviations to form syndrome variant indicator variables. Veterans positive for any of the syndrome variants met the criteria for the GWI Research case definition. The case definition was validated in a sample of Gulf War veterans from a Veterans Affairs medical center[32] and in the full national sample of the USMHS.[7] This case definition, reproduced in new samples of veterans by applying the factor weights to the same battery of symptoms questions, was used in the initial neurological, neuroimaging, and genetic studies to identify abnormalities of autonomic function and brain processing as well as genetic predisposition of GWI.[14,15,17,22,33]

Subsequently, Fukuda et al. developed the U.S. Centers for Disease Control and Prevention (CDC) case definition, from a questionnaire of 10 typical GWI symptoms requiring at least 1 symptom from at least 2 of 3 symptom domains to be considered a case.[3] Later, Steele developed the Kansas case definition from a questionnaire of 37 symptoms, requiring at least 3 symptoms representing at least 3 of 6 symptom domains.[4] In addition, veterans with any of 10 comorbid conditions were excluded as cases. However, as the veterans aged, the exclusions eliminated too many valid cases; therefore, in recent studies the number of comorbidities excluded has been reduced or eliminated.[34,35] The symptoms comprising the CDC and Kansas definitions overlapped 90% with those used for the GWI Research definition. Because the groups meeting the GWI Research definition are generally a close subset of the larger numbers meeting the CDC and Kansas definitions, the GWI Research definition represents a consensus of the three definitional approaches (Table S1).

Prevalence Case–Control Subsample

In a second stage of sampling (Figure 1), we obtained a prevalence case–control subsample from the 8,020 USMHS survey participants. Near the end of the CATI interview questionnaire, all participants meeting either the GWI Research or Kansas case definitions and a random sample of all others were invited to participate in the second stage of the study involving collection of a blood sample (Table 1). Of the 2,103 individuals from whom we could obtain a blood sample, were deployed to the KTO, and were not in the earlier study from which the GWI Research case definition was developed (Figure 1), we selected as the cases all veterans meeting the GWI Research case definition[2] ($n = 508$). The controls comprised an independent sample of those meeting none of the case definitions ($n = 508$). These selection criteria

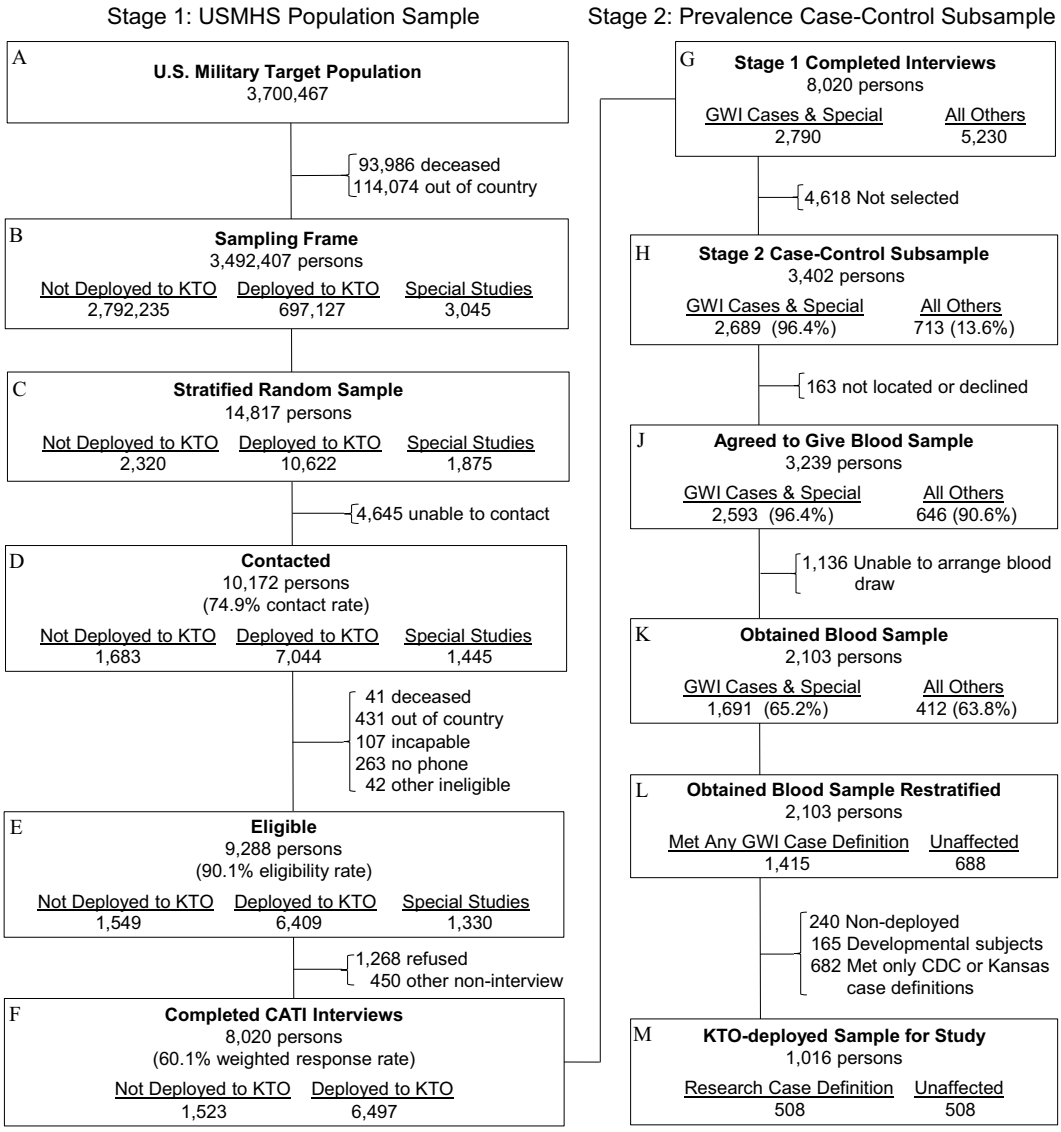

Figure 1. Selection of the stage 1 USMHS population sample and the stage 2 prevalence case–control subsample. (A) Includes all U.S. military personnel on active duty or in the Reserves or National Guard on 2 August 1990. (B) Those not deployed to the Kuwaiti Theater of Operations (KTO) included medically nondeployable personnel. "Special Studies" included twins, members of the 24th Reserve Naval Mobile Construction Battalion (Seabees), and parents of children with Goldenhar Complex. Counts for subgroups are suppressed to maintain confidentiality according to terms of the Certificate of Confidentiality. (C) The sampling frame was stratified by age, sex, race, service branch, military rank, active duty/reserve status, special studies strata, and KTO location on 20 January 1991.[7] (D) The denominator of the survey-weighted contact rate includes the number of known survey-eligible persons and the estimated number of eligible persons among those with an undetermined survey eligibility status. (E) The eligibility rate was calculated with survey weights applied among sample members with known survey eligibility. (F) The survey-weighted response rate is the American Association for Public Opinion Research's Response Rate 4 (RR4), calculated as the number of confirmed and estimated eligible cases among those initially selected for the CATI phase of the study.[31] (G) The universe for the current prevalence case–control study includes all subjects in (F). (H) In selecting the prevalence case–control sample, the CATI algorithm selected all GWI cases by the Research and Kansas case definitions and a 12.5% random sample of the rest of the USMHS participants. The slight deviations from these selection percentages resulted from late adjustments in the CATI algorithm. (L) Veterans who met any GWI case definition specifically met the GWI Research, CDC, or Kansas (without exclusions) case definitions. The 165 developmental subjects excluded after (L) were those from the Seabees battalion who participated in the initial study to develop the GWI Research case definition. (M) To minimize misclassification, the 508 who met the GWI Research case definition—a close subset of the CDC and Modified Kansas definitions—were separated from those meeting only the CDC or Modified Kansas definitions and were used as cases in the prevalence case-control study. Unaffected subjects, those meeting none of the three GWI case definitions, constituted the controls in the prevalence case–control study. The left-hand column (A–F) of the figure was adapted from the original USMHS sampling flowchart published in Iannacchione et al.[7] Note: CATI, computer-assisted telephone interview; CDC, U.S. Centers for Disease Control and Prevention; GWI, Gulf War Illness; USMHS, U.S. Military Health Survey.

were used to minimize misclassification in the case and control groups. The protocol was approved by the institutional review boards of University of Texas (UT) Southwestern Medical Center, RTI International, the U.S. Army, and the Department of Veterans Affairs; all subjects gave verbal informed consent for the survey and written informed consent for phlebotomy and genetic testing.

Blood Collection

From 2009 to 2010, blood samples were drawn by licensed phlebotomists in or near the subjects' homes. Heparinized whole blood and serum separator tubes were cooled immediately and shipped overnight in a refrigeration pack to the program's laboratory, where serum and plasma were aliquoted and leukocytes

Table 1. Distributions of demographic and environmental exposure characteristics in the GWI cases and controls ($n = 1{,}016$).

	GWI cases ($n = 508$)		Controls ($n = 508$)	
	Frequency	Percent	Frequency	Percent
Age (y)				
<49	294	57.9	356	70.0
≥49	213	41.9	152	30.0
Missing	1	0.2	0	0.0
Sex				
Male	396	78.0	477	93.9
Female	112	22.0	31	6.1
Service branch				
Navy	78	15.4	184	36.22
Army	338	66.5	182	35.8
Marines	63	12.4	90	17.7
Air Force	29	5.7	52	10.3
Force status				
Guard/Reserve	166	32.7	188	37.0
Active duty	342	67.3	320	63.0
Military rank				
Officer	28	5.5	80	15.7
Enlisted	480	94.5	428	84.3
Combat exposure scale				
Light	273	53.7	422	83.1
Light to moderate	102	20.1	52	10.2
Moderate to heavy	60	11.8	10	2.0
Missing	73	14.4	24	4.7
Special strata[a]				
Yes	67	13.2	286	56.3
No	441	86.8	222	43.7
Heard nerve agent alarms				
No	111	21.9	287	56.5
Yes	397	78.2	221	43.5
Unit located in Khamisiyah plume				
No	354	69.7	386	76.0
Yes	120	23.6	56	11.0
Missing (unit location unknown)	34	6.7	66	13.0

Note: GWI, Gulf War Illness.
[a]Special studies strata included twin pairs, members of the 24th Reserve Naval Mobile Construction Battalion (Seabees), and parents of children with Goldenhar Complex. Counts for subgroups are suppressed to maintain confidentiality according to terms of the Certificate of Confidentiality.[7]

processed for DNA, all of which were frozen at −80°C for later assay. Blood samples received warm were rejected and redrawn.

Measures of Environmental Exposure

On the basis of reports of widespread sounding of nerve agent alarms and identifications of chemical warfare agents in U.S. and Coalition troop positions from 18 to 19 January 1991 shortly after Coalition bombing of Iraqi chemical weapon production and storage facilities early in the air campaign[36–38] (Figure 2), we asked the following survey question in the USMHS CATI interviews as the measure of low-level nerve agent exposure: "During the time period from 2 August 1990 to 31 July 1991, did the alarms on the chemical warfare detection devices in areas where you were living or working ever go off while you were present there?" To provide a measure of the dose of the exposure, we included this follow-up question requesting a numerical response: "Approximately how many different times did the chemical nerve gas detection devices go off in your area?" In a previous analysis,[39] we established defined categories, 0, 1, 2–9, and ≥10 d, to provide sufficient within-category sample sizes. Each individual's KTO location on 20 January 1991 was included among the USMHS stratification variables to maximize the power of the sample for testing the effects of this measure on illness.

To test an alternative hypothesis on nerve agent exposure during the conflict period, we obtained from the U.S. Army Center for Unit Records Research (CURR) three binary indicator variables for each member of the USMHS sample indicating whether their military unit was located in the plume of possible nerve agent exposure on days 1, 2, or 3 after the demolition of the Khamisiyah ammunition dump 10 d after the end of the conflict period, 10–12 March 1991. The area of the plume had been estimated in 2000 by computerized atmospheric transport and dispersion (ATD) modeling.[40] We created a dichotomous measure indicating Khamisiyah plume exposure on any days and a trichotomous measure of exposure coded 0 d, 1 d, or 2 d (no veteran in our sample was exposed for all 3 d).

Genotyping

Genomic DNA was purified from subjects' leukocytes by the UT Southwestern's Genomics and Microarray Core. High-throughput genotyping was performed on a Bio-Rad iQ5 Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.), using the Taqman SNP genotyping/allelic discrimination assay per the manufacturer's protocol (Thermo Fisher). Labeled proprietary primers and probes for the *PON1* Q192R (SNP ID rs662) and butyrylcholinesterase K variant (alanine/threonine A539T; rs1803274) genotyping assays were designed, manufactured, and validated by the manufacturer. All samples were run in triplicate, and each plate included positive and negative control samples.[25]

Serum Esterase Assays

All aliquoting, transfer and dilution of serum, addition of substrates, and enzymatic analysis were performed on a robotic BioTek system equipped with a Precision XS automated sample pipetting station, Twister II microplate handler, and a Synergy HT microplate spectrophotometer as previously described.[25] Serum samples were thawed at room temperature, vortexed for 20 s, centrifuged at $10{,}000 \times g$ for 2 min at 4°C and supernatants diluted with 5 mM Tris-HCl, pH 7.4 buffer containing 1 mM $CaCl_2$ immediately before analysis. Serum hydrolytic activity for the following PON1-defining substrates was measured in a 96-well plate format according to published methods: paraoxonase,[41] diazoxonase,[41] and arylesterase.[22] All samples were run in triplicate and rerun if coefficients of variation (CVs) were higher than 15%; a quality control serum sample was run with all samples to ensure accuracy; and every batch of substrate reagents was tested for purity against a standard. The overall interassay CVs from the repeat testing of the quality control serum sample in every batch were 4.9%, 4.0%, and 3.4% for arylesterase, diazoxonase, and paraoxonase activities, respectively, and the intra-assay CVs among the three triplicate runs of each sample were 5.2%, 2.7%, and 1.4%.

Subjects were classified into the three phenotypes of the Q192R polymorphism by the association of their diazoxonase and paraoxonase activity levels, and subjects' level of serum hydrolytic activity of the two isoenzymes of the Q192R polymorphism, the PON1 type Q and R isoenzymes, was quantified as described in Figure 3. A more detailed description of the genotyping and enzymatic assay methods has been published.[25]

Serum butyrylcholinesterase activity was measured using 50 μM benzoylcholine (Sigma-Aldrich) in a protocol modified from published methods[42,43] for a reaction volume of 200 μL. Benzoylcholine assays were conducted in ultraviolet-transparent microplates (Greiner Bio); the reactions were monitored using absorbance at 240 nm at 25°C. Serum dilutions and reactions were made in 0.067 M Na/K phosphate buffer pH 7.4. Inhibition of activity by dibucaine was used to identify the "atypical" (A) phenotype.[44] Dibucaine number (DN) is the percentage inhibition of activity caused by 10 μM of dibucaine.

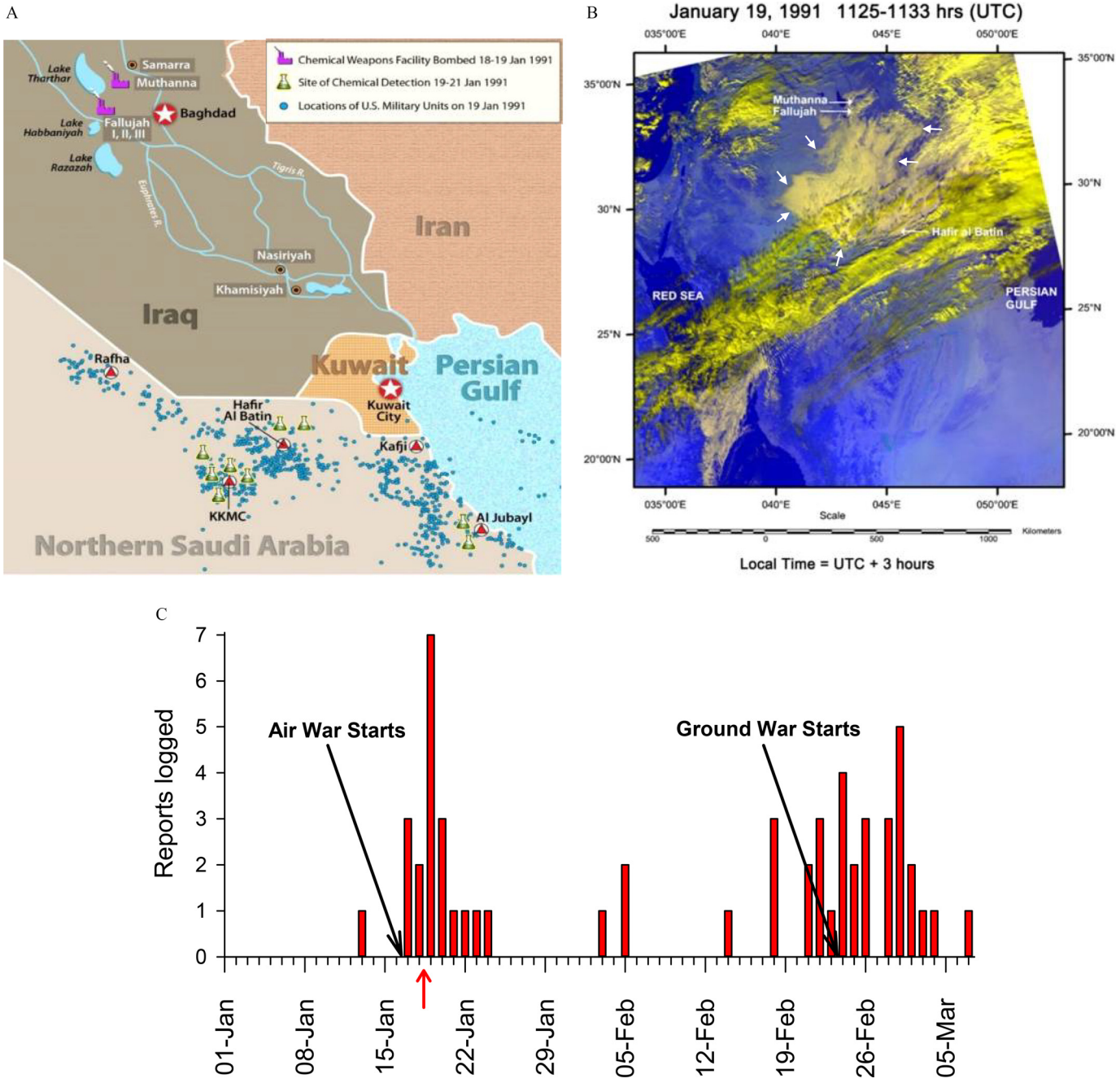

Figure 2. (A) Map of the Kuwaiti Theater of Operations (KTO) showing the locations of major chemical weapons storage facilities bombed on the night of 18–19 January and location of U.S. military units and sites of sarin and other chemical weapon detections on 19–21 January. (B) Weather satellite image of the large debris cloud (light tan in color and demarcated by white arrows) containing dispersed chemical weapon vapor.[38] Sequential images from every-2-h passes showed the debris rising from the bombed chemical weapons storage facilities at Muthanna and Fallujah and drifting southward to encompass U.S. troop positions. This image was taken at approximately 14:30 h local time on 19 January showing the debris reaching Hafir Al Batin, the day 10,000 nerve agent alarms began sounding and chemical weapons experts using sophisticated equipment detected ambient sarin and other agents at multiple sites across U.S. positions.[37,87] The light green cloud bank extending from northeast to southwest indicated a stationary weather front that held the sarin-containing debris cloud over U.S. troop positions for a week. (C) Numbers of reports of alarms, warnings, etc., logged within the Nuclear, Biological, and Chemical cells of the Central Command, Army Central Command, and VII Army Corps during the Conflict Period of the Gulf War[36]; the red vertical arrow marks the night of 18–19 January just before the satellite image in (B) was taken. Figures (A) and (B) reproduced from *Neuroepidemiology*[38] by permission of S. Karger AG, Basel, and (C) from the June 1994 report of the Defense Science Board Task Force on Persian Gulf War Health Effects.[36]

Statistical Methods

Descriptive analyses. Analyses to estimate the population prevalence of GWI and the main environmental risk factors from the entire deployed sample of the USMHS ($n = 6,497$) were weighted with the survey adjustment weights to correct for the unequal probabilities of selection from the strata and selection biases from inability to locate and refusal to participate.[7,45] Correct standard errors, allowing for the complex survey design and the weights, were performed with SAS survey procedures (version 9.4; SAS Institute) and SUDAAN software (RTI International).[46]

Because the information was obtained either from personal interviews or from official military files, few values were missing. One subject's missing age was imputed to 45, the mean age

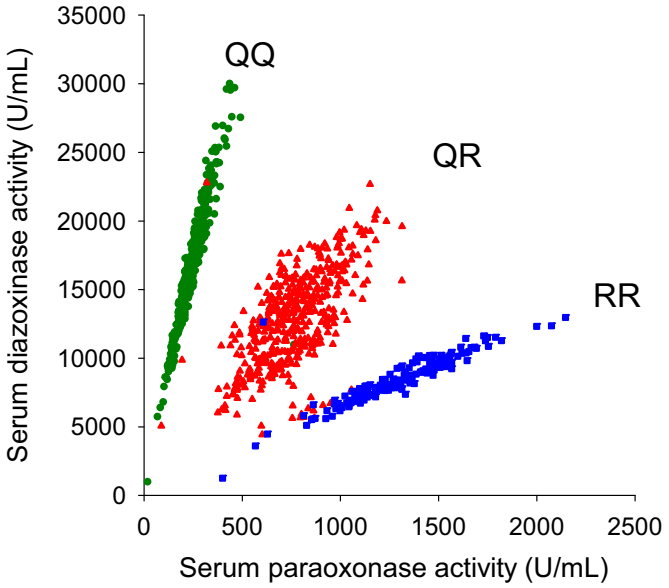

Figure 3. Classification of the 1,016 Gulf War-era veterans of the prevalence case-control sample into the three PON1 Q192R phenotypes (QQ, QR, and RR) by their serum hydrolytic activity for substrates diazoxon ("diazoxonase activity") and paraoxon ("paraoxonase activity"). QQ subjects have only the Q isoenzyme (green circles); RR subjects have only the R isoenzyme (blue squares); and the QR subjects have some of each (red triangles). The relative amounts of the Q and R isoenzymes in each QR subject is estimated by an interpolation equation.[22,25]

of the sample. The "moderate to heavy," and "heavy" categories of the Combat Exposure Scale were combined with the "moderate" category because of small numbers, and in multivariable modeling, the categories were represented as dummy variables with "light" as the referent and "missing data" as a separate dummy variable. In analyses involving possible exposure to the Khamisiyah plume, 9.5% of subjects were omitted because their unit locations were unknown.

Test of monotonicity of GWI-risk factor associations. We tested for monotonicity (i.e., a consistently increasing or decreasing function) of the associations of GWI with ordinal measures of the environmental and genetic risk factors by calculating Stuart and Kendall's *tau*-c (τ_c) nonparametric correlation coefficient and its asymptotic standard error and the *p*-value from the Cochran-Armitage nonparametric trend test.[47] τ_c is distributed from −1 to 1 and estimates the strength of monotonicity, and the trend test provides its statistical significance. For the graphical plots for Figure 4, we analyzed these ordinal categorical risk factors as dummy variables in logistic regression models to generate odds ratios (ORs) appropriate for case–control studies.

Test of gene–environment independence. We tested for gene–environment independence with a logistic regression model in the 508 controls to test the association of having the *PON1* R allele with having heard nerve agent alarms, controlling for the potential confounders.[48]

GxE interaction analyses. We tested our prestated hypothesis on the association of GWI with the GxE interaction of the *PON1* gene and low-level nerve agent exposure in the Gulf War. The effect of the *PON1* gene was represented by the Q192R polymorphism where veterans with the minor RR genotype were the high-risk group and by the serum activity level of the Q isoenzyme, categorized at quartiles—both measured from the blood samples collected in the second stage blood study of the USMHS. We analyzed two binary measures of the environmental exposure: *a*) the USMHS CATI question asking whether the subject recalled having heard nerve agent alarms sounding in their area, and *b*) the indicator variables of veterans' having been in the computer-modeled plume from the Khamisiyah demolitions.

The analyses were carried out to conform with the recommendations of Knol and VanderWeele[49] for displaying the results of interactions in genetic epidemiological studies in the familiar 4×2 table with a single reference category, which extended the earlier Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations.[50] We provided the final measures of interaction on both the additive and multiplicative scales with and without adjustment for confounding. We used Rothman's synergy index (S) for the additive-scale measure to facilitate comparison with the multiplicative-scale measure represented by the prevalence OR (POR) of the interaction term from logistic regression, both of which range from 0 to plus infinity with equivalency point of 1.0. We also provided the aRERI for the most extreme comparison (highest vs. lowest risk categories), which is used to define the level of interaction.[51]

We calculated adjusted S and aRERI and their 95% confidence intervals (CIs) with Zou's SAS macro,[52] which we modified by adding the front end of the Li and Chambless macro[53] to automate the interface with the logistic regression output and adding a *p*-value calculation for the synergy index. Zou's asymmetric CIs are more accurate than the symmetric ones of Hosmer and Lemeshow[54] and, in our large samples, gave virtually the same additive scale results as Richardson and Kaufman's linear OR approach[55] (Table S2).

Analyses of confounding. In the design of the USMHS, we reviewed past studies to identify the most likely important confounding variables and selected age, sex, service branch, military rank, active-duty status, special strata, and Keane Combat Exposure Scale[56] (Table 1). Anticipating their potential importance, we obtained these variables from the official military personnel file and incorporated them as stratification variables in the design of the USMHS survey[7] to provide sufficient statistical power to control for them in the analyses.

Sensitivity analyses of the effect of recall bias. To assess the effect of recall bias[30] on our estimates of the GxE interaction, following the approach of Greenland and Lash,[57] we stipulated the likely ranges of the sensitivity (Se) and specificity (Sp) of veterans' recollection of hearing nerve agent alarms expected under recall bias. Assuming recall bias would result in higher sensitivity and lower specificity with the two parameters varying inversely in the cases and the reverse in the controls, we examined the range of values of Se from 0.90 to 0.99 and Sp of 0.90 to 0.70 in the GWI cases, in comparison with values of Se 0.80 to 0.90 and Sp 0.95 to 0.90 in the unaffected controls. These values of Se and Sp for recalling nerve agent alarms are in a higher range than expected for recall of usual life experiences because soldiers were trained to recognize the characteristic sound of the M8A1 organophosphate detectors in their camps and immediately don the potentially life-saving impervious rubber suits, masks, and gloves—highly memorable events. We assumed that GWI would not affect recall of wartime events because the veterans' subjective memory complaints have been attributed to deficits of attention, concentration, and working memory (i.e., executive function) and affective issues[58] that developed weeks to months after the war, when long-term memories had already been fixed.

Using a SAS macro we developed and validated for the sensitivity analysis (Tables S3 and S4), we repeatedly recalculated the strength of the GxE interaction on both the additive and multiplicative scales after correcting the distribution of the environmental exposure variable for values across the range of Se and Sp in the case and control groups expected from recall bias.

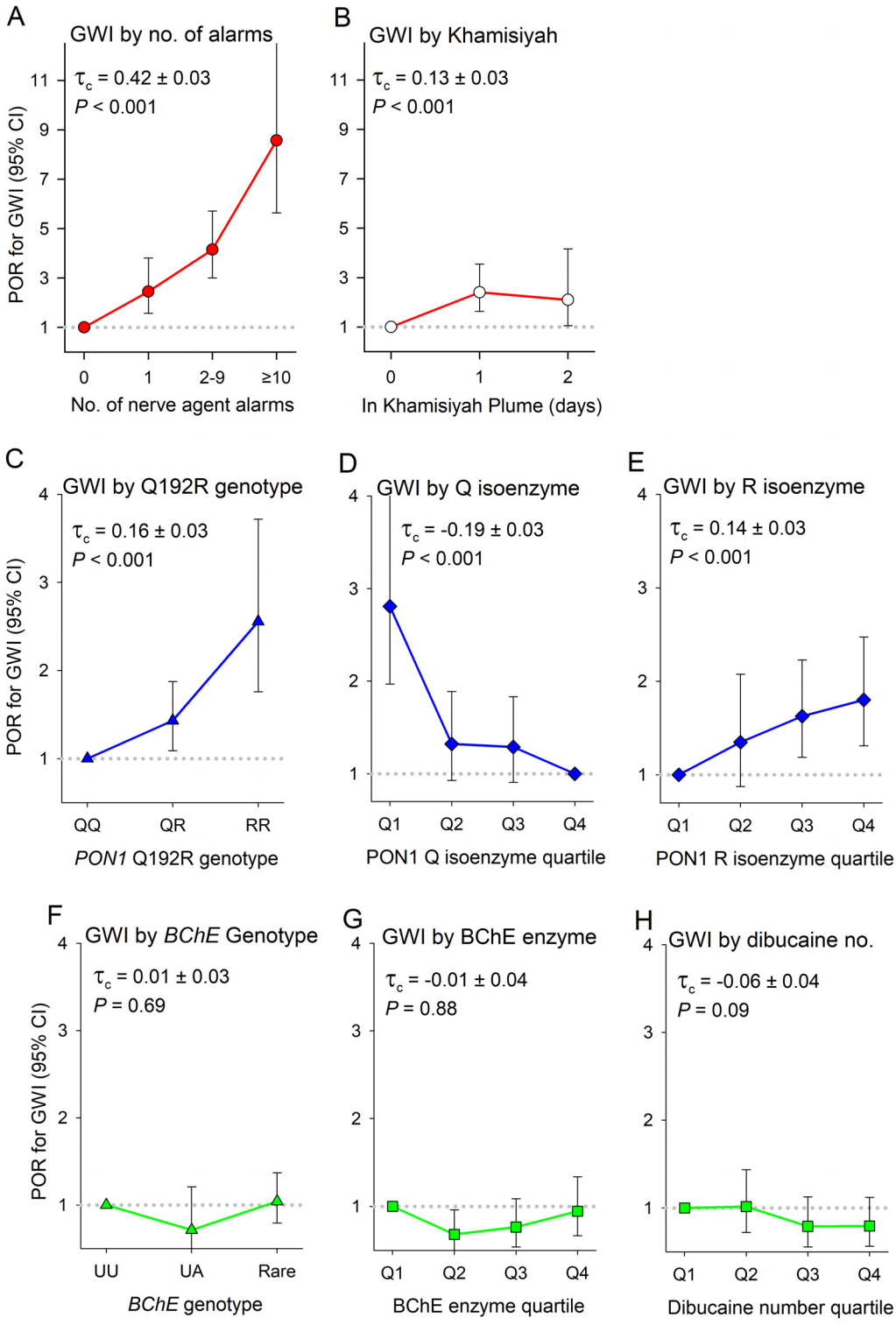

Figure 4. The association of measures of low-level nerve agent exposure and genetic predisposition with Gulf War Illness. The measures are (A) the number of times veterans were possibly exposed to low-level nerve agent indicated by nerve agent alarms sounding where they were present; (B) the number of days veterans were in an area exposed to a possible plume of low-level nerve agent from postwar demolition of artillery shells containing sarin and cyclosarin in the Khamisiyah ammunition dump according to U.S. government computer modeling; (C) the *PON1* Q192R genotype; (D) serum PON1 192Q isoenzyme activity; (E) serum PON1 192R isoenzyme activity; (F) the *BChE* genotype; (G) serum BChE enzyme activity; and (H) dibucaine number. Statistics are unadjusted. Note: BChE, serum butyrylcholinesterase activity; CI, confidence interval; GWI, Gulf War Illness; POR, prevalence odds ratio; τ_c, Stuart and Kendall's tau-c nonparametric correlation coefficient and its asymptotic standard error testing monotonicity; P, two-tailed significance test of $\tau_c = 0$. The numerical values for the graphs are given in Table S5.

Table 2. Interaction on the additive and multiplicative scales of hearing nerve agent alarms and *PON1* Q192R genotype on GWI.

	PON1 Q192R genotype						PORs for *PON1* Q192R genotypes within strata of alarms	
	QQ		QR		RR			
	Cases/ controls (*n*)	POR (95% CI)	Cases/ controls (*n*)	POR (95% CI)	Cases/ controls (*n*)	POR (95% CI)	QR vs. QQ	RR vs. QQ
Heard nerve agent alarms								
No	43/130	1.0	50/120	1.26 (0.78, 2.03)[a] $p=0.34$	18/37	1.47 (0.76, 2.85)[a] $p=0.25$	1.26 (0.78, 2.03) $p=0.34$	1.47 (0.76, 2.85) $p=0.25$
Yes	129/104	3.75 (2.44, 5.77)[a] $p<0.001$	177/96	5.57 (3.64, 8.53)[a] $p<0.001$	91/21	13.10 (7.29, 23.55)[a] $p<0.001$	1.49 (1.04, 2.13) $p=0.03$	3.49 (2.04, 6.00) $p<0.001$
POR (95% CI) for alarms within strata of genotypes		3.75 (2.44, 5.77)[b] $p<0.001$		4.43 (2.93, 6.69)[b] $p<0.001$		8.91 (4.27, 18.60)[b] $p<0.001$		
Additive scale: Synergy index (95% CI)								
Unadjusted		1.0		1.52 (0.93, 2.48) $p=0.09$		3.76 (1.91, 7.37) $p<0.001$		
Adjusted for confounders		1.0		1.87 (0.95, 3.67) $p=0.07$		4.71 (1.82, 12.19)[c] $p=0.001$		
Multiplicative scale: POR (95% CI) from LR interaction term								
Unadjusted		1.0		1.18 (0.65, 2.14) $p=0.59$		2.38 (1.01, 5.57) $p=0.047$		
Adjusted for confounders		1.0		1.45 (0.70, 2.97) $p=0.32$		3.41 (1.20, 9.72) $p=0.02$		

Note: The synergy index is a measure of interaction on the additive scale; it has the same distribution as the POR, viz., 0 to plus infinity with 1.0 as the equivalency point indicating no association. The ratio of the PORs, obtained from the interaction term in a logistic regression analysis, is a measure of interaction on the multiplicative scale. The potential confounders controlled for in the adjusted models include: age (years), sex (M, F), service branch [Army (referent), Navy, Air Force, Marines], rank (officer, enlisted), active duty vs Guard/ Reserve, special strata (yes, no), Combat Exposure Scale [0 = missing, 1 = light (referent), 2 = light to moderate, 3 = moderate to heavy and heavy]. One subject's missing age was imputed to the mean age of the sample. The analyses included 508 cases and 508 controls. aRERI, relative excess risk due to interaction adjusted for measured confounding; CI, confidence interval; GWI, Gulf War Illness; LR, logistic regression; PON1, paraoxonase-1; POR, prevalence odds ratio.
[a]Each of these odds ratio is derived from the 2x2 table formed by the cases and controls of this cell and those in the QQ-No alarm cell as the referent.
[b]Each of these odds ratio is derived from the 2x2 table formed by the cases and controls of the 2 cells above it with the top cell as the referent.
[c]This value of the adjusted synergy index corresponds to an aRERI of 7.69 (95% CI: 2.71, 19.13).

Sensitivity analysis for unmeasured confounding. To assess the likelihood that confounding by unmeasured characteristics might explain away the main findings on the GxE interaction, we performed sensitivity analyses for interaction on both the additive and multiplicative scales under unmeasured confounding by the methods of VanderWeele et al.[59] Given that the *PON1* Q192R polymorphism was unconfounded (randomly assigned at birth and unknown to participants), the analysis estimated the bias in the GxE interaction from confounding for varying levels of association of the unmeasured confounder (U) with both hearing nerve agent alarms (stipulated prevalence of U in those who did and those who did not hear alarms) and GWI (stipulated rate ratio of U with GWI). We validated our SAS macros by comparing the interaction parameters generated by entering 1.0 for the three design parameters with values in Table 2 generated by our other macros.

Throughout the study, $p \le 0.05$ was the threshold for statistical significance.

Results

Characteristics of the Study Population and the Prevalence Case–Control Sample

By applying the survey weights to our random sample of the 1991 U.S. military population, we estimated that 91,910 (13.6%; 95% CI: 11.8, 15.4) of the approximately 677,000 personnel deployed to the KTO for the 1991 Persian Gulf War met our GWI Research case definition; 286,359 (42.3%; 95% CI: 39.9, 44.7) personnel reported hearing nerve agent alarms; and 95,131 (14.1%; 95% CI: 13.2, 15.0) were located in the hypothesized Khamisiyah plume.

The frequency distributions of the baseline characteristics of the 1,016 veterans randomly selected from the survey participants for the USMHS second stage prevalence case–control sample are given in Table 1. In comparison with the controls ($n = 508$), the cases ($n = 508$) were older, more likely female, Army, active duty, enlisted, and combat exposed.

Association of GWI with Environmental Risk Factors

Logistic regression analysis found that GWI increased monotonically over the number of nerve agent alarms heard to a crude POR of 8.57 (95% CI: 5.64, 12.04) for those who heard 10 or more alarms, and the τ_c and trend tests strongly supported a monotonic association (Figure 4A; Table S5). GWI was weakly associated with the number of days of exposure to the Khamisiyah plume to a crude POR of 2.10 (95% CI: 1.06, 4.16), although monotonicity was supported (Figure 4B).

Association of GWI with Genotype Distributions

In comparison with the Q192R genotype distribution of the control veterans (QQ 47.7%, QR 39.3%, and RR 13.0%), the GWI group was significantly enriched in the RR genotype (QQ 24.2%, QR 57.7%, and RR 18.1%; $p < 0.001$). The distribution of *PON1* Q192R genotypes in the control group did not differ significantly from the Hardy-Weinberg equilibrium ($p = 0.54$), but that of the GWI cases did ($p = 0.04$). GWI increased monotonically across the three *PON1* Q192R genotypes (Figure 4C).

In contrast, there was no significant difference in the *BChE* genotype distributions between the controls (UU 60.2%, UK 36.0%, and KK 3.8%) and the veterans with GWI (UU 58.6%, UK 39.4%, and KK 2.0%; $p = 0.82$) (Figure 4F).

Association of GWI with Phenotype Distributions

GWI decreased monotonically with increasing PON1 type Q isoenzyme levels and increased monotonically with increasing R isoenzyme levels (Figure 4D,E), whereas GWI was not significantly associated with the BChE enzyme activity level or the dibucaine number (Figure 4G,H).

Table 3. Interaction on the additive and multiplicative scales of hearing nerve agent alarms and PON1 type Q isoenzyme level on GWI.

	PON1 type Q isoenzyme activity level (quartiles)								POR (95% CI) for PON-Q quartiles Within strata of alarms		
	4th Quartile (lowest risk)		3rd Quartile (mid-low risk)		2nd Quartile (mid-high risk)		1st Quartile (highest risk)				
	Cases/ controls (*n*)	POR (95% CI)	Cases/ controls (*n*)	POR (95% CI)	Cases/ controls (*n*)	POR (95% CI)	Cases/ controls (*n*)	POR (95% CI)	3rd vs. 4th quartile	2nd vs. 4th quartile	1st vs. 4th quartile
Heard nerve Agent alarms											
No	29/83	1.0	29/74	1.12 (0.61, 2.05)[a] $p=0.71$	25/77	0.93 (0.45, 1.72)[a] $p=0.82$	28/53	1.51 (0.81, 2.82)[a] $p=0.19$	1.12 (0.61, 2.05) $p=0.71$	0.93 (0.50, 1.72) $p=0.82$	1.51 (0.81, 2.82) $p=0.19$
Yes	74/70	3.03 (1.77, 5.16)[a] $p<0.001$	89/62	4.10 (2.41, 7.00)[a] $p<0.001$	88/50	5.04 (2.92, 8.71)[a] $p=0.001$	146/39	10.71 (6.18, 18.59)[a] $p<0.001$	1.36 (0.86, 2.15) $p=0.19$	1.67 (1.03, 2.68) $p=0.04$	3.54 (2.19, 5.73) $p<0.001$
PORs (95% CI) for alarms within strata of PON-Q activity		3.03 (1.77, 5.16)[b] $p<0.001$		3.66 (2.14, 6.27)[b] $p<0.001$		5.42 (3.73, 9.58)[b] $p<0.001$		7.09 (3.97, 12.64)[b] $p<0.001$			
Additive scale: Synergy index (95% CI)											
Unadjusted		1.0		1.45 (0.71, 2.96) $p=0.31$		2.07 (0.95, 4.47) $p=0.07$		3.83 (1.94, 7.55) $p<0.001$			
Adjusted for confounders		1.0		1.38 (0.57, 3.35) $p=0.48$		2.48 (0.96, 6.39) $p=0.06$		3.89 (1.60, 9.49)[c] $p=0.003$			
Multiplicative scale: POR (95% CI) from LR interaction term											
Unadjusted		1.0		1.21 (0.57, 2.58) $p=0.62$		1.79 (0.82, 3.91) $p=0.14$		2.34 (1.07, 5.15) $p=0.034$			
Adjusted for confounders		1.0		1.07 (0.43, 2.68) $p=0.88$		2.30 (0.90, 5.89) $p=0.08$		2.78 (1.08, 7.17) $p=0.03$			

Note: The synergy index is a measure of interaction on the additive scale; it has the same distribution as the POR, viz., 0 to plus infinity with 1.0 as the equivalency point indicating no association. The ratio of the PORs, obtained from the interaction term in a logistic regression analysis, is a measure of interaction on the multiplicative scale. The potential confounders controlled for in the adjusted models include: age (years), sex (M, F), service branch [Army (referent), Navy, Air Force, Marines], rank (officer, enlisted), active duty vs. Guard/Reserve, special strata (yes, no), Combat Exposure Scale [0 = missing, 1 = light (referent), 2 = light to moderate, 3 = moderate to heavy and heavy]. One subject's missing age was imputed to the mean age of the sample. The analyses included 508 cases and 508 controls. Comparable tables for the PON1 R isoenzyme, diazoxonase, arylesterase, paraoxonase, and BChE enzyme are given in Tables S8–S15. aRERI, relative excess risk due to interaction adjusted for measured confounding; BChE, butyrylcholinesterase; CI, confidence interval; GWI, Gulf War Illness; LR, logistic regression; POR, prevalence odds ratio; PON1, paraoxonase-1; POR, prevalence odds ratio.

[a]Each of these odds ratio is derived from the 2x2 table formed by the cases and controls of this cell and those in the QQ-No alarm cell as the referent.

[b]Each of these odds ratio is derived from the 2x2 table formed by the cases and controls of the 2 cells above it with the top cell as the referent.

[c]This value of the adjusted synergy index corresponds to an aRERI of 5.91 (95% CI: 2.49, 13.45).

Table 4. Sensitivity analysis of the effect of differential misclassification of the environmental variable (hearing nerve agent alarms) on the association of GWI with the GxE interaction between the PON1 RR vs. QQ genotype and having heard nerve agent alarms on the additive and multiplicative scales.

		Interaction on the additive scale[a]				
		Controls				
Cases		Se: 1.00	0.90	0.85	0.80	0.80
Se	Sp	Sp: 1.00	0.90	0.90	0.90	0.95
1.00	1.00	3.76 (1.91, 7.37)[b]	— —	— —	— —	— —
0.90	0.90	—	4.45 (2.35, 8.41)	4.57 (2.40, 8.68)	4.74 (2.46, 9.14)	4.73 (2.40, 9.32)
0.90	0.80	—	4.70 (2.43, 9.10)	4.86 (2.48, 9.52)	5.09 (2.53, 10.23)	5.13 (2.46, 10.73)
0.90	0.70	—	5.10 (2.53, 10.29)	5.34 (2.58, 11.03)	5.69 (2.63, 12.32)	5.85 (2.51, 13.62)
0.95	0.80	—	4.55 (2.25, 9.19)	4.70 (2.27, 9.72)	4.92 (2.28, 10.65)	4.93 (2.14, 11.37)
0.95	0.70	—	4.90 (2.29, 10.48)	5.12 (2.29, 11.41)	5.45 (2.27, 13.11)	5.55 (2.07, 14.91)

		Interaction on the multiplicative scale[c]				
		Controls				
Cases		Se: 1.00	0.90	0.85	0.80	0.80
Se	Sp	Sp: 1.00	0.90	0.90	0.90	0.95
1.00	1.00	2.38 (1.01, 5.57)[b]	— —	— —	— —	— —
0.90	0.90	—	4.04 (1.50, 10.90)	4.12 (1.53, 11.05)	4.21 (1.58, 11.26)	4.05 (1.52, ,10.78)
0.90	0.80	—	4.12 (1.58, 10.75)	4.20 (1.62, 10.90)	4.30 (1.67, 11.10)	4.14 (1.61, 10.62)
0.90	0.70	—	4.25 (1.69, 10.68)	4.33 (1.73, 10.83)	4.43 (1.78, 11.03)	4.26 (1.72, 10.55)
0.95	0.80	—	3.11 (1.31, 7.41)	3.17 (1.34, 7.51)	3.25 (1.38, 7.64)	3.12 (1.33, 7.31)
0.95	0.70	—	3.20 (1.38, 7.46)	3.26 (1.41, 7.56)	3.34 (1.45, 7.69)	3.21 (1.40, 7.35)

Note: —, no data; GxE, gene-environment interaction; GWI, Gulf War illness; Se, sensitivity; Sp, specificity.
[a]Cells of the upper table contain the unadjusted synergy index (95% CI).
[b]From Table 2.
[c]Cells of the lower table contain the unadjusted prevalence odds ratio (95% CI) of the interaction term from logistic regression.

Independence of Genotype and Environmental Exposure

In the 508 controls, the adjusted POR for having the R allele and answering affirmatively to the survey question on hearing nerve agent alarms—controlling for age, sex, service branch, military rank, active duty/reserve status, special strata, and combat exposure—was 1.18 (95% CI: 0.81, 1.73, $p = 0.35$).

GxE Interaction of PON1 Q192R Genotype and Nerve Agent Exposure

GWI was significantly associated with a strong monotonically increasing, synergistic GxE interaction between environmental exposure to nerve agent alarms and having the *PON1* R allele, as tested on the additive scale by the synergy index. Specifically, in comparison with those having no R allele (the QQ homozygotes), those with one R allele (QR heterozygotes) had a synergy index of 1.87 (95% CI: 0.95, 3.67), and those with two R alleles (RR homozygotes) 4.71 (95% CI: 1.82, 12.19) (Table 2). When tested on the multiplicative scale, the POR of the interaction term in the logistic regression analysis showed the same monotonic increase in GWI over the number of R alleles (Table 2). The aRERI, adjusted for confounding, was 7.69 (95% CI: 2.71, 19.13) (Table 2).

Using the Khamisiyah plume as the indicator of environmental exposure, the tests for GxE interaction yielded similarly elevated associations on both additive and multiplicative scales, but the synergy values were small and imprecise (Tables 1; Table S6).

GxE Interaction of PON1 Q Isoenzyme Activity and Nerve Agent Exposure

GWI was significantly associated with a monotonically increasing, synergistic GxE interaction, as indicated on the additive scale by the synergy index, with environmental exposure to nerve agent alarms and having the given type Q isoenzyme activity level. Specifically, in comparison with the fourth quartile of type Q activity (the most protected), those in the third quartile had a synergy index of 1.38 (95% CI: 0.57, 3.35); those in the second quartile, 2.48 (95% CI: 0.96, 6.39); and those in the first quartile (the least protected), 3.89 (95% CI: 1.60, 9.49) (Table 3). Similar results were found with the analysis on the multiplicative scale (Table 3). The aRERI, adjusted for confounding, was 5.91 (95% CI: 2.49, 13.45) (Table 3).

The tests for GxE interaction involving the Khamisiyah plume yielded similar elevations that were likewise small and imprecise (Tables 1; Table S7).

Tests for GxE Interaction Involving BChE and Additional PON1-Related Enzyme Activities

Whereas all *BChE* variant genotypes had lower BChE serum activity than the normal U/U genotype (Figure S1), we found no significant GxE interaction for GWI between either the *BChE* genotype or the BChE serum activity level with having heard nerve agent alarms or having taken pyridostigmine (Tables S8–S11).

We found no statistically significant GxE interactions involving the PON1 R isoenzyme or paraoxonase activity (mediated mostly by the R isoenzyme) or arylesterase activity (mediated by both Q and R isoenzymes), but we did find a significant GxE interaction involving diazoxonase activity (mediated mostly by the Q isoenzyme) (Figure 3; Tables S12–S15).

Is the GxE Interaction Due to Recall Bias?

Correcting the unadjusted values of the GxE interaction for the combinations of Se and Sp of recalling nerve agent alarms expected under recall bias increased the association of the GxE interaction with GWI above the baseline uncorrected values on both the additive ($S = 3.76$) and multiplicative (POR = 2.38) scales, thus forcing them further away from the null (Table 4).

Confounding by Measured and Unmeasured Characteristics

Adjustment for the seven measured confounding variables generally increased the point estimates of the GxE interaction, correcting negative confounding on both the additive and multiplicative scales (Tables 2 and 3). Correcting for all combinations of the plausible ranges of association of an unmeasured confounder with the independent and dependent variables failed to reduce the aRERI to unity, and only the most prevalent unmeasured confounders and those most strongly associated with the independent and dependent variables forced the lower confidence limit down to encompass unity (Table S16). Similarly, the point estimate of the GxE interaction on the multiplicative scale was resistant to all but the most extremely prevalent and strongly associated unmeasured confounders, whereas its confidence interval included unity into the plausible range (Table S17).

Discussion

Our study arrived at several findings that address the possible causal role of low-level nerve agent exposure in GWI. We found strong associations between GWI and both hearing nerve agent alarms and having *PON1* Q192R genotypes that would increase susceptibility to injury from nerve agent. Moreover, there was a strong GxE interaction by which the strength of the association of GWI with hearing nerve agent alarms increased monotonically with the subject's number of *PON1* 192R alleles. This monotonic interaction remained strong when substituting the phenotypic PON1 Q isoenzyme activity level for the *PON1* genotype. Our sensitivity analysis found that misclassification due to recall bias in Gulf War veterans' self-reporting of hearing nerve agent alarms appears to have forced the test for GxE interactions toward the null, tending to obscure the association rather than generating a falsely positive one. Finally, we found that the *PON1* genotype and hearing nerve agent alarms were independent and the findings robust to both measured and unmeasured confounding, supporting a mechanistic GxE interaction.

Plausibility of a Link between Low-Level Nerve Agent Exposure and GWI

Our first finding was a monotonically increasing association of GWI with categories of the number of times nerve agent alarms sounded where the veterans were working or living. The plausibility of this finding was supported by several other findings. Although the presence of nerve agent during the Gulf War was initially doubted, subsequent evidence substantiated widespread low-level exposure among U.S. troops from long-distance transit of fallout from the bombing of Iraqi chemical weapon research, production, and storage facilities west of Baghdad early in the 1991 Gulf War,[36–38] summarized in Figure 2. GWI has been statistically associated with various self-report measures and plume model indicators of nerve agent exposure in most of the epidemiological studies of GWI in which it has been analyzed[8,21,39,60–64] (Table S18). Putative nerve agent exposure has also been linked epidemiologically with death from brain cancer in Gulf War veterans in the first decade after the war[65] as well as clinically with abnormalities of veterans' neurocognitive function,[60,63] regional brain volumes,[66–69] and brain white matter integrity.[70] A similar body of epidemiologic and clinical evidence has demonstrated a neurocognitive syndrome and dysfunction of the central nervous system and postural instability similar to those in GWI in the Japanese survivors of sarin attacks (by the Aum Shinrikyo cult on a housing subdivision in Matsumoto, Japan, in 1994 and on the Tokyo subway in 1995) in whom the severity of the chronic illness increased with the documented dose of sublethal sarin exposure.[71–73]

Plausibility of Protection from Nerve Agent Effects by the Q Isoenzyme

Our second finding was a monotonically increasing association between GWI and the *PON1* Q192R genotype, with the risk increasing with the number of R alleles. This association was shown to be physiologically important by our phenotypic measurements of the serum enzymatic products, the Q and R isoenzymes, of the corresponding alleles. Given that the Q isoenzyme hydrolyzes chemical warfare nerve agents like sarin more efficiently than the R isoenzyme does,[23] the Q isoenzyme would be expected to provide greater protection from adverse effects of nerve agent exposure. We found that the serum activity level of the PON1 type Q isoenzyme, but not the type R isoenzyme, was inversely associated with, i.e., protective from, GWI. Thus, the biased distribution of the genotype appears to reflect enrichment of the GWI group with veterans having the R allele, which would have made them more susceptible to neurotoxicity from low-level sarin nerve agent.

In contrast, we found, as in two previous studies,[22,27] no association of GWI with low serum BChE activity, rare *BChE* genetic variants, or a GxE interaction involving either of them and low-level nerve agent exposure.

Evidence of Effect Modification by PON1 Q192R Genotype and Q Isoenzyme

If the associations of GWI with nerve agent exposure and with the *PON1* Q192R polymorphism were indicating a causal relationship, we would predict that the genotype would act as an effect modifier for the GWI–nerve agent exposure association. Our third finding was confirmation of this GxE interaction. Specifically, the strength of the association between the dichotomous measures of GWI and hearing nerve agent alarms during the 1991 Gulf War conflict period increased monotonically over the genotype categories of the number of R alleles and inversely over the categories of type Q isoenzyme serum activity. Moreover, the change in the combined effect from one category to the next was significantly greater than the sum of the independent effects of the environmental exposure and the genotype, indicating a synergistic interaction as measured on both the additive and multiplicative scales.

In contrast to the strong association of GWI with low-level nerve agent exposure during the conflict period, indicated by hearing nerve agent alarms, the evidence provided only weak support for a GxE interaction with low-level nerve agent exposure from postwar demolition of Iraqi artillery shells containing nerve agent in an ammunition dump near Khamisiyah, hypothesized by the computer-generated plume model. The evidence for nerve agent exposure during the conflict period, including transit of a likely

nerve agent–containing debris cloud from bombed Iraqi storage sites, large numbers of nerve agent alarms, credible detections of nerve agent by chemical weapons experts, and widespread concern about nerve agent exposures at the time, was far stronger than after the Khamisiyah ammunition dump demolition, which was not suspected until years after the war.[37–40] The weaker support for a GxE interaction involving the Khamisiyah plume may have been due to relatively lower amounts of nerve agent released, low statistical power due to the lower prevalence of personnel exposed, or exposure misclassification by the plume model, which has been criticized.[74]

Because the ability to hydrolyze sarin (sarinase activity) is largely specific to the PON1 Q isoenzyme activity,[23] as predicted we found no GxE interaction with PON1 R isoenzyme activity or with PON1's paraoxonase activity, which is highly dependent on the R isoenzyme. Subjects' serum activity levels of the Q and R isoenzymes tend to be inversely associated (Figure 3). The Q isoenzyme efficiently hydrolyzes many organophosphate nerve agents, whereas the R isoenzyme does not.[23,75] Subjects' level of serum diazoxonase activity (vertical axis in Figure 3) is correlated with their level of Q isoenzyme activity and thus is associated with protection from nerve agents. Subjects' serum paraoxonase activity (horizontal axis in Figure 3) is correlated with their R isoenzyme activity and thus is weakly associated with protection from nerve agents. Phenylacetate ("arylesterase activity") is equally associated with Q and R activity and thus is only moderately associated with protection from nerve agents.

Effect of Recall Bias on the GxE Interaction

A potential limitation of our study is that recall bias in veterans' reporting of having heard nerve agent alarms might have inflated its association with GWI—a charge that, without evidence, has in the past contributed to the etiologic contribution of nerve agent to GWI being disregarded.[76] Finding a GxE interaction, however, provided a unique opportunity to determine the likely magnitude and direction of this potential bias. Using the well-described epidemiological approach to this problem,[57] we recalculated the GxE interactions on the additive and multiplicative scales over ranges of the sensitivity and specificity of recall bias, assuming that the recall of ill veterans would display relatively low specificity and high sensitivity, because ill individuals tend to recall exposures more vividly and possibly embellish, whereas well controls would display high specificity but low sensitivity, because healthy individuals tend to recall less vividly.[77] Given that RT-PCR performed in triplicate measured the *PON1* genotypes with negligible error and that the *PON1* gene is known to have a causal effect in modifying organophosphate effects,[78] the results showed that correcting for the examined ranges of recall accuracy forced the GxE interactions further from the null, indicating that misclassification of the environmental variable by recall bias had actually tended to obscure the association rather than manufacturing a false one, as long incorrectly assumed. This result was to be expected, because when the environmental and genetic variables are independent and the associations are unconfounded[79]—both conditions supported by our analysis—measurement error in the environmental variable always biases the GxE interaction toward the null, and conversely, under these assumptions, a significant GxE interaction cannot be due to misclassification of the environmental variable.[48]

Strength of Causal Inference Suggested by the Level of GxE Interaction

The epidemiological literature recognizes three levels of GxE interaction distinguished by the strength of evidence for a causal relationship: *a*) statistical interaction, *b*) mechanistic interaction, and *c*) biological, or functional, interaction.[51] A statistical interaction is typically one demonstrated by a statistically significant coefficient on the interaction term in a logistic regression analysis. Known in the epidemiology literature as an interaction on the multiplicative scale, it is an unreliable indicator of a causal interaction at the physiologic level, because statistical interactions frequently result from reasons other than causal interactions.

A mechanistic interaction must satisfy more rigorous conditions that lead to a stronger inference of causality. The central requirement for a mechanistic interaction is the demonstration that at least some individuals (though not necessarily all) will have the disease if both environmental and genetic exposures are present but not if just one of them is.[51] Known as Rothman's criteria for "sufficient cause"[80,81] or "compositional epistasis,"[51,82] this requirement can be established by a statistically significant interaction on the additive scale, conditioned on monotonicity of the associations of the outcome with both the environmental and genetic variables. Monotonicity means that the risk functions of the exposures with the disease are always either increasing or neutral but never decreasing for any individual; that is, an exposure that increases the disease risk is never also preventive. In addition, the multivariable models demonstrating the additive interaction must control for confounding of the associations of the outcome with both the environmental exposure and the genetic variable, and the gene and environmental variables must occur independently of each other.[79]

Our analysis met all of these conditions: controlling for potential confounders for the association of GWI with both the number of nerve agent alarms heard and the *PON1* Q192R genotype; demonstrating gene and environmental variable independence; and establishing monotonicity of GWI prevalence rate with both the environmental and genetic variables. For our purposes, the monotonicity of nerve agent exposure is certain: Not only was the GWI a strongly increasing monotonic function of the number of nerve agent alarms heard (Figure 4A), but the large body of published evidence has identified no circumstances in which sarin is preventive of neurotoxic brain effects.[83] Whereas warnings from nerve agent alarms were designed to prevent high-level, potentially fatal exposures from nerve gas attacks, which did not occur in the Gulf War, they do not prevent low-level exposure from atmospheric fallout as occurred in the Gulf War. The alarms' level of detection was set high enough that by the time they sounded (or after they stopped), nearby soldiers had already been (or continued to be) exposed to low levels of nerve agent long enough to produce lasting neurological damage demonstrable by electroencephalogram[38,84] (Figure S2).

Likewise, monotonicity is certain for the genetic variables *PON1* Q192R genotype and type Q isoenzyme activity. GWI was a monotonically increasing function of the number of R alleles and a monotonically decreasing function of the quartiles of type Q isoenzyme activity (Figure 4C,D), and the Q isoenzyme's only known effect on sarin is to inactivate it by hydrolysis and thus mitigate its neurotoxic effects.[23]

Given, then, that both the environmental and genetic variables satisfy the monotonicity requirement, an unconfounded aRERI >0 would establish mechanistic interaction.[51] Our analysis found that, controlling for potential confounders, aRERI $= 7.69$ (95% CI: 2.71, 19.13) for the interaction with the number of R alleles, and aRERI $= 5.91$ (95% CI: 2.49, 13.45) for interaction with the serum level of type Q isoenzyme activity. Because the lower 95% confidence limits of both analyses exceeded the aRERI >0 criterion, these findings establish a mechanistic interaction.

We are then left with the question whether this mechanistic GxE interaction actually represents a biological, or functional,

interaction, i.e., where low-level nerve agent exposure and the genotype of the *PON1* Q192R polymorphism interacted biochemically to produce GWI. Although the presence of a mechanistic interaction suggests this, it must be further supported by a convincing mosaic of basic research establishing a compatible biological mechanism.[51] Extensive biochemical and clinical research has elucidated both the adverse brain effects of low-level sarin at doses comparable to levels to which soldiers would have been exposed during the Gulf War (Table S19) and the ability of the *PON1* gene through its type Q isoenzyme to efficiently hydrolyze sarin at the low physiological concentrations expected from subsymptomatic exposure and prevent its neurotoxic effects[23] (Table S20).

Confounding by Measured Characteristics

In the design of the USMHS we measured all the characteristics considered possible confounders and incorporated them in the random sampling design as stratification variables to ensure adequate power to control for them in the multivariable analyses.[7] Controlling for these seven confounders only strengthened the GxE interaction effects, indicating negative confounding.

Confounding by Unmeasured Characteristics

Our finding a significant GxE interaction under a strong assumption of GxE independence means that either the GxE interaction truly exists or it is due to an interaction between the genotype and an unmeasured confounder of the environment effect.[59] In our case, however, the fact that the *PON1* Q192R genotype was randomly assigned at birth and was unknown to participants during and after the war makes associations between the genotype and unmeasured confounders unlikely.[59] If there is no genotype–confounder association, then the GxE interaction on either the additive or multiplicative scales cannot be biased by confounding even if the unmeasured confounder is not controlled for.[51] In the unlikely event that, notwithstanding these considerations, unmeasured confounding was present, our sensitivity analysis found that our measures of GxE interaction on the additive and multiplicative scales were robust to all plausible patterns and ranges of unmeasured confounding. In addition, the finding of negative confounding by the measured confounders limits the possibility of confounding by unmeasured characteristics associated with those we measured.

Three possible confounders not analyzed in this study are pesticide exposure,[85] rocket or jet fuel or vehicle exhaust exposure,[86] and severe fright from hearing nerve agent alarms.[8] Both organophosphate pesticides and fuel or exhaust fumes are capable of causing falsely positive nerve agent alarms from the M8A1 nerve agent detectors,[86] but exposures to both were ubiquitous long before the approximately 10,000 alarms began sounding at the start of the air campaign when Coalition bombing of Iraqi chemical weapon facilities released the fallout cloud that reached U.S. troop concentrations just as sarin was detected at multiple sites[37,87,88,39,38] (Figure 2). Whereas heavy repetitive exposure to organophosphate pesticides might cause chronic cognitive problems, the PON1 R isoenzyme is the more efficient detoxifier of most pesticides.[23] Jet exhaust is not known to have neurotoxic effects,[89] and PON1 does not metabolize petroleum-derived hydrocarbons. Thus, if the wide array of pesticides used in the Gulf War[85] or jet exhaust were responsible for most of the alarms, the GxE interaction involving the Q isoenzyme would not have occurred. Although severe fright can produce posttraumatic stress disorder (PTSD), psychological explanations including PTSD do not explain GWI fully,[11] and the studies suggesting that GWI was PTSD were shown to be falsely positive misinterpretations of psychological screening tests.[90]

Limitations of the Study

A limitation of this study is that it focused on environmental exposures to low-level sarin nerve agent to the exclusion of other risk factors implicated in epidemiological studies, such as exposures to pesticides, pyridostigmine bromide, antibiotics, immunizations, insect repellants, and psychological effects of deployment.[91] This focus was due to the unique opportunity afforded by the *PON1* Q192R polymorphism to develop objective genetic and biochemical support for their etiologic role in GWI by studying GxE interactions. Another limitation was that we were unable to apply Mendelian randomization to strengthen the evidence for a causal relationship between low-level nerve agent exposure and GWI, because the *PON1* Q192R polymorphism is acting as an effect modifier rather than an instrumental variable, which is required to apply that approach.[92] The USMHS survey participation rate of 60% leaves the possibility of selection bias; however, the application of survey weights that control for such bias had little effect on parameter estimates.[7] Our decision not to adjust significance levels for multiple statistical testing is justified by use of prespecified hypotheses, the strong biochemical basis for the findings, and the relationships among the various exposures studied.

Conclusion

Our study supports the prestated hypothesis of a GxE interaction between the *PON1* Q192R polymorphism and low-level nerve agent exposure measured by recall of hearing nerve agent alarms during the Gulf War conflict. Contemporaneous weather satellite images[38] have removed the objection that originally discounted the role of nerve agent in GWI; selection of our GWI cases and controls from a large, representative sample of U.S. Gulf War–era veterans avoided selection bias; potential confounding by both measured and unmeasured confounders was ruled out; and sensitivity testing suggested that recall bias would have obscured the GxE interaction (biased toward the null), rather than causing a false positive one. The GxE interaction met the rigorous criteria for a mechanistic interaction, constituting a higher level of evidence for a causal link than a mere statistical epidemiological association. The prior research linking low-level sarin exposure with brain pathology compatible with GWI and demonstrating the biochemically modifying effects of the PON1 Q isoenzyme on the effects of sarin satisfy the higher standard of evidence for a biological interaction. These findings constitute strong evidence for a causal role of low-level nerve agent in GWI.

Acknowledgments

A large research team of survey specialists at RTI International contributed importantly to the design and performed the field work for the national CATI survey of Gulf War–era veterans. Research leaders included K.A. Considine, V.G. Iannacchione, J.A. Dever, C.P. Carson, H. Best, C. Bann, D. Creel, B. Alexander, A. Lewis-Evans, L. Trofimovich, K. Pate, A. Kenyon, J. Morton, C. Hill, and R.E. Mason. UT Southwestern team members who contributed substantially included C. Garcia, A. Lamb, R. Thompson, E. Cordell, J. Escobar, and W. Marshall.

This work was funded by the following: U.S. Army Medical Research and Materiel Command grant number DAMD17-01-1-0741; Department of Veterans Affairs Medical Center, Dallas, Texas, IDIQ contract VA549-P-0027; and the Office of the Assistant Secretary of Defense for Health Affairs, through the Gulf War Illness Research Program under Award No. W81XWH-16-1-0740. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not

necessarily endorsed by the U.S. Departments of Defense or Veterans Affairs.

Human subjects protection review: The protocol was approved by the institutional review boards of UT Southwestern Medical Center, RTI International, the U.S. Army, and the Department of Veterans Affairs; all subjects gave verbal informed consent for the survey and written informed consent for phlebotomy and genetic testing.

Data availability: The data that support the findings of this study are the property of the U.S. Department of Veterans Affairs and were analyzed by the authors under a data use agreement with the University of Texas Southwestern Medical Center.

Role of the funding source: None of the funding agencies had any role in the design and conduct of the study; in collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript. The corresponding author had full access to all the data (including statistical reports and tables) in the study, takes responsibility for the integrity of the data and the accuracy of the data analysis, and takes final responsibility for the decision to submit for publication.

References

1. Institute of Medicine (US) Committee on Measuring the Health of Persian Gulf Veterans. 1999. *Gulf War Veterans: Measuring Health*. Hernandez LM, Durch JS, Blazer DG II, Hoverman IV, eds. Washington, DC: National Academies Press. PMID: 25077231, https://doi.org/10.17226/9636.
2. Haley RW, Kurt TL, Hom J. 1997. Is there a Gulf war syndrome? Searching for syndromes by factor analysis of symptoms. JAMA 277(3):215–222, PMID: 9005271, https://doi.org/10.1001/jama.1997.03540270041025.
3. Fukuda K, Nisenbaum R, Stewart G, Thompson WW, Robin L, Washko RM, et al. 1998. Chronic multisymptom illness affecting air force veterans of the Gulf war. J Am Med Assoc 280(11):981–988, PMID: 9749480, https://doi.org/10.1001/jama.280.11.981.
4. Steele L. 2000. Prevalence and patterns of Gulf war illness in Kansas veterans: association of symptoms with characteristics of person, place, and time of military service. Am J Epidemiol 152(10):992–1002, PMID: 11092441, https://doi.org/10.1093/aje/152.10.992.
5. Institute of Medicine. 2010. *Gulf War and Health, Vol. 8: Update of Health Effects of Serving in the Gulf War*. Washington, DC: National Academies Press. PMID: 25032373, https://doi.org/10.17226/12835.
6. Kang HK, Li B, Mahan CM, Eisen SA, Engel CC. 2009. Health of US veterans of 1991 Gulf War: a follow-up survey in 10 years. J Occup Environ Med 51(4):401–410, PMID: 19322107, https://doi.org/10.1097/JOM.0b013e3181a2feeb.
7. Iannacchione VG, Dever JA, Bann CM, Considine KA, Creel D, Carson CP, et al. 2011. Validation of a research case definition of Gulf War illness in the 1991 US military population. Neuroepidemiology 37(2):129–140, PMID: 21986258, https://doi.org/10.1159/000331478.
8. Nisenbaum R, Barrett DH, Reyes M, Reeves WC. 2000. Deployment stressors and a chronic multisymptom illness among Gulf War veterans. J Nerv Ment Dis 188:259–266, PMID: 10830562, https://doi.org/10.1097/00005053-200005000-00002.
9. Cherry N, Creed F, Silman A, Dunn G, Baxter D, Smedley J, et al. 2001. Health and exposures of United Kingdom Gulf war veterans. Part II: the relation of health to exposure. Occup Environ Med 58(5):299–306, PMID: 11303078, https://doi.org/10.1136/oem.58.5.299.
10. Gray GC, Reed RJ, Kaiser KS, Smith TC, Gastañaga VM. 2002. Self-reported symptoms and medical conditions among 11,868 Gulf War-era veterans: the Seabee Health Study. Am J Epidemiol 155(11):1033–1044, PMID: 12034582, https://doi.org/10.1093/aje/155.11.1033.
11. Ismail K, Kent K, Brugha T, Hotopf M, Hull L, Seed P, et al. 2002. The mental health of UK Gulf war veterans: phase 2 of a two phase cohort study. BMJ 325(7364):576–579, PMID: 12228134, https://doi.org/10.1136/bmj.325.7364.576.
12. Haley RW, Marshall WW, McDonald GG, Daugherty MA, Petty F, Fleckenstein JL. 2000. Brain abnormalities in Gulf War syndrome: evaluation with 1H MR spectroscopy. Radiology 215(3):807–817, PMID: 10831703, https://doi.org/10.1148/radiology.215.3.r00jn48807.
13. Menon PM, Nasrallah HA, Reeves RR, Ali JA. 2004. Hippocampal dysfunction in Gulf War Syndrome. A proton MR spectroscopy study. Brain Res 1009(1–2):189–194, PMID: 15120596, https://doi.org/10.1016/j.brainres.2004.02.063.
14. Haley RW, Spence JS, Carmack PS, Gunst RF, Schucany WR, Petty F, et al. 2009. Abnormal brain response to cholinergic challenge in chronic encephalopathy from the 1991 Gulf War. Psychiatry Res 171(3):207–220, PMID: 19230625, https://doi.org/10.1016/j.pscychresns.2008.05.004.
15. Gopinath K, Gandhi P, Goyal A, Jiang L, Fang Y, Ouyang L, et al. 2012. FMRI reveals abnormal central processing of sensory and pain stimuli in ill Gulf War veterans. Neurotoxicology 33(3):261–271, PMID: 22327017, https://doi.org/10.1016/j.neuro.2012.01.014.
16. Haley RW, Charuvastra E, Shell WE, Buhner DM, Marshall WW, Biggs MM, et al. 2013. Cholinergic autonomic dysfunction in veterans with Gulf War illness: confirmation in a population-based sample. JAMA Neurol 70(2):191–200, PMID: 23407784, https://doi.org/10.1001/jamaneurol.2013.596.
17. Moffett K, Crosson B, Spence JS, Case K, Levy I, Gopinath K, et al. 2015. Word-finding impairment in veterans of the 1991 Persian Gulf War. Brain Cogn 98:65–73, PMID: 26114921, https://doi.org/10.1016/j.bandc.2015.05.005.
18. Turner MP, Hubbard NA, Himes LM, Faghihahmadabadi S, Hutchison JL, Bennett IJ, et al. 2016. Cognitive slowing in Gulf War Illness predicts executive network hyperconnectivity: study in a population-representative sample. Neuroimage Clin 12:535–541, PMID: 27672557, https://doi.org/10.1016/j.nicl.2016.08.022.
19. Tillman GD, Calley CS, Buhl VI, Chiang HS, Haley RW, Hart J, Jr., et al. 2017. Electrophysiological correlates of semantic memory retrieval in Gulf War Syndrome 2 patients. J Neurol Sci 373:66–72, PMID: 28131230, https://doi.org/10.1016/j.jns.2016.12.023.
20. Tillman GD, Spence JS, Briggs RW, Haley RW, Hart J, Jr., Kraut MA. 2019. Gulf War illness associated with abnormal auditory P1 event-related potential: evidence of impaired cholinergic processing replicated in a national sample. Psychiatry Res Neuroimaging 283:7–15, PMID: 30453127, https://doi.org/10.1016/j.pscychresns.2018.11.006.
21. Haley RW, Kurt TL. 1997. Self-reported exposure to neurotoxic chemical combinations in the Gulf War. A cross-sectional epidemiologic study. JAMA 277(3):231–237, PMID: 9005273, https://doi.org/10.1001/jama.1997.03540270057027.
22. Haley RW, Billecke S, La Du BN. 1999. Association of low PON1 type Q (type A) arylesterase activity with neurologic symptom complexes in Gulf war veterans. Toxicol Appl Pharmacol 157(3):227–233, PMID: 10373407, https://doi.org/10.1006/taap.1999.8703.
23. La Du BN, Billecke S, Hsu C, Haley RW, Broomfield CA. 2001. Serum paraoxonase (PON1) isozymes: the quantitative analysis of isozymes affecting individual sensitivity to environmental chemicals. Drug Metab Dispos 29(4 Pt 2):566–569, PMID: 11259353.
24. Costa LG, Cole TB, Vitalone A, Furlong CE. 2005. Measurement of paraoxonase (PON1) status as a potential biomarker of susceptibility to organophosphate toxicity. Clin Chim Acta 352(1–2):37–47, PMID: 15653099, https://doi.org/10.1016/j.cccn.2004.09.019.
25. Teiber JF, Kramer GL, Haley RW. 2013. Methods for measuring serum activity levels of the 192 Q and R isoenzymes of paraoxonase 1 in QR heterozygous individuals. Clin Chem 59(8):1251–1259, PMID: 23894162, https://doi.org/10.1373/clinchem.2012.199331.
26. Loewenstein-Lichtenstein Y, Schwarz M, Glick D, Norgaard-Pedersen B, Zakut H, Soreq H. 1995. Genetic predisposition to adverse consequences of anticholinesterases in 'atypical' BChE carriers. Nat Med 1(10):1082–1085, PMID: 7489367, https://doi.org/10.1038/nm1095-1082.
27. Steele L, Lockridge O, Gerkovich MM, Cook MR, Sastre A. 2015. Butyrylcholinesterase genotype and enzyme activity in relation to Gulf War illness: preliminary evidence of gene-exposure interaction from a case-control study of 1991 Gulf War veterans. Environ Health 14:4, PMID: 25575675, https://doi.org/10.1186/1476-069X-14-4.
28. Michalovicz LT, Kelly KA, Sullivan K, O'Callaghan JP. 2020. Acetylcholinesterase inhibitor exposures as an initiating factor in the development of Gulf War Illness, a chronic neuroimmune disorder in deployed veterans. Neuropharmacology 171:108073, PMID: 32247728, https://doi.org/10.1016/j.neuropharm.2020.108073.
29. National Academies of Sciences, Engineering, and Medicine. 2016. *Gulf War and Health: Volume 10: Update of Health Effects of Serving in the Gulf War, 2016*. Cory-Slechta D, Wedge R, eds. Washington, DC: National Academies Press. https://doi.org/10.17226/21840.
30. Glass DC, Sim MR. 2006. The challenges of exposure assessment in health studies of Gulf War veterans. Philos Trans R Soc Lond B Biol Sci 361(1468):627–637, PMID: 16687267, https://doi.org/10.1098/rstb.2006.1822.
31. American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 8th ed. Chicago, IL: American Association for Public Opinion Research. https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.
32. Haley RW, Luk GD, Petty F. 2001. Use of structural equation modeling to test the construct validity of a case definition of Gulf War syndrome: invariance over developmental and validation samples, service branches and publicity. Psychiatry Res 102(2):175–200, PMID: 11408056, https://doi.org/10.1016/S0165-1781(01)00241-4.
33. Tillman GD, Calley CS, Green TA, Buhl VI, Biggs MM, Spence JS, et al. 2012. Event-related potential patterns associated with hyperarousal in Gulf War

Illness syndrome groups. Neurotoxicology 33(5):1096–1105, PMID: 22691951, https://doi.org/10.1016/j.neuro.2012.06.001.
34. Brewer KL, Mainhart A, Meggs WJ. 2018. Double-blinded placebo-controlled cross-over pilot trial of naltrexone to treat Gulf War Illness. Fatigue Biomed Health Behav 6(3):132–140, https://doi.org/10.1080/21641846.2018.1477034.
35. Dursa E, Barth S, Porter B, Schneiderman A. 2018. Gulf War Illness in the 1991 Gulf War and Gulf era veteran population: an application of the Centers for Disease Control and Prevention and Kansas case definitions to historical data. J Milit Vet Health 26:43–50. https://jmvh.org/wp-content/uploads/2018/05/Dursa-Original-Article.pdf [accessed 11 April 2022].
36. Defense Science Board. 1994. Report of the Defense Science Board Task Force on Persian Gulf War Health Effects. Washington, DC: Office of the Under Secretary of Defense for Acquisition and Technology. https://gulflink.health.mil/dsbrpt/ [accessed 19 February 2022].
37. U.S. Senate Committee on Banking, Housing, and Urban Affairs. 1994. The Riegle Report: U.S. Chemical and Biological Warfare-Related Dual Use Exports to Iraq and Their Possible Impact on the Health Consequences of the Persian Gulf War. http://people.cryst.bbk.ac.uk/~toxin/riegle/riegle1.html [accessed 19 February 2022].
38. Tuite JJ, Haley RW. 2013. Meteorological and intelligence evidence of long-distance transit of chemical weapons fallout from bombing early in the 1991 Persian Gulf war. Neuroepidemiology 40(3):160–177, PMID: 23257977, https://doi.org/10.1159/000345123.
39. Haley RW, Tuite JJ. 2013. Epidemiologic evidence of health effects from long-distance transit of chemical weapons fallout from bombing early in the 1991 Persian Gulf war. Neuroepidemiology 40(3):178–189, PMID: 23258108, https://doi.org/10.1159/000345124.
40. Office of the Special Assistant for Gulf War Illnesses. 2000. Case narrative: US demolition operations at Khamisiyah. https://gulflink.health.mil/khamisiyah_ii/index.html [accessed 11 April 2022].
41. Richter RJ, Furlong CE. 1999. Determination of paraoxonase (PON1) status requires more than genotyping. Pharmacogenetics 9(6):745–753, PMID: 10634137, https://doi.org/10.1097/00008571-199912000-00009.
42. Kalow W, Lindsay HA. 1955. A comparison of optical and manometric methods for the assay of human serum cholinesterase. Can J Biochem Physiol 33(4):568–574, PMID: 13240528, https://doi.org/10.1139/y55-071.
43. Lockridge O. 1990. Genetic variants of human serum cholinesterase influence metabolism of the muscle relaxant succinylcholine. Pharmacol Ther 47(1):35–60, PMID: 2195556, https://doi.org/10.1016/0163-7258(90)90044-3.
44. Kalow W, Genest K. 1957. A method for the detection of atypical forms of human serum cholinesterase; determination of dibucaine numbers. Can J Biochem Physiol 35(6):339–346, PMID: 13437188, https://doi.org/10.1139/y57-041.
45. Valliant R, Dever JA. 2018. *Survey Weights: A Step-by-Step Guide to Calculation*. College Station, TX: Stata Press.
46. LaVange LM, Stearns SC, Lafata JE, Koch GG, Shah BV. 1996. Innovative strategies using SUDAAN for analysis of health surveys with complex samples. Stat Methods Med Res 5:311–329, PMID: 8931198, https://doi.org/10.1177/096228029600500306 [accessed 11 April 2022].
47. Agresti A. 2010. *Analysis of Ordinal Categorical Data*. 2 ed. New York, NY: John Wiley & Sons.
48. García-Closas M, Thompson WD, Robins JM. 1998. Differential misclassification and the assessment of gene-environment interactions in case-control studies. Am J Epidemiol 147(5):426–433, PMID: 9525528, https://doi.org/10.1093/oxfordjournals.aje.a009467.
49. Knol MJ, VanderWeele TJ. 2012. Recommendations for presenting analyses of effect modification and interaction. Int J Epidemiol 41(2):514–520, PMID: 22253321, https://doi.org/10.1093/ije/dyr218.
50. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. 2007. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology 18(6):805–835, PMID: 18049195, https://doi.org/10.1097/EDE.0b013e3181577511.
51. VanderWeele TJ, Knol MJ. 2014. A tutorial on interaction. Epidemiol Methods 3(1):33–72. https://doi.org/10.1515/em-2013-0005.
52. Zou GY. 2008. On the estimation of additive interaction by use of the four-by-two table and beyond. Am J Epidemiol 168(2):212–224, PMID: 18511428, https://doi.org/10.1093/aje/kwn104.
53. Li R, Chambless L. 2007. Test for additive interaction in proportional hazards models. Ann Epidemiol 17(3):227–236, PMID: 17320789, https://doi.org/10.1016/j.annepidem.2006.10.009.
54. Hosmer DW, Lemeshow S. 1992. Confidence interval estimation of interaction. Epidemiology 3(5):452–456, PMID: 1391139, https://doi.org/10.1097/00001648-199209000-00012.
55. Richardson DB, Kaufman JS. 2009. Estimation of the relative excess risk due to interaction and associated confidence bounds. Am J Epidemiol 169(6):756–760, PMID: 19211620, https://doi.org/10.1093/aje/kwn411.
56. Keane TM, Fairbank JA, Caddell JM, Zimering RT, Taylor KL, Mora CA. 1989. Clinical evaluation of a measure to assess combat exposure. Psychol Assess 1(1):53–55, https://doi.org/10.1037/1040-3590.1.1.53.
57. Greenland S, Lash TL. 2008. Bias analysis. In: *Modern Epidemiology*. Rothman KJ, Greenland S, Lash TL, eds. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 352–357.
58. Jeffrey MG, Krengel M, Kibler JL, Zundel C, Klimas NG, Sullivan K, et al. 2019. Neuropsychological findings in Gulf War Illness: a review. Front Psychol 10:2088, PMID: 31616335, https://doi.org/10.3389/fpsyg.2019.02088.
59. Vanderweele TJ, Mukherjee B, Chen J. 2012. Sensitivity analysis for interactions under unmeasured confounding. Stat Med 31(22):2552–2564, PMID: 21976358, https://doi.org/10.1002/sim.4354.
60. White RF, Proctor SP, Heeren T, Wolfe J, Krengel M, Vasterling J, et al. 2001. Neuropsychological function in Gulf War veterans: relationships to self-reported toxicant exposures. Am J Ind Med 40(1):42–54, PMID: 11439396, https://doi.org/10.1002/ajim.1070.
61. Kang HK, Mahan CM, Lee KY, Murphy FM, Simmens SJ, Young HA, et al. 2002. Evidence for a deployment-related Gulf War syndrome by factor analysis. Arch Environ Health 57(1):61–68, PMID: 12071362, https://doi.org/10.1080/00039890209602918.
62. Lindem K, Heeren T, White RF, Proctor SP, Krengel M, Vasterling J, et al. 2003. Neuropsychological performance in Gulf War era veterans: traumatic stress symptomatology and exposure to chemical-biological warfare agents. J Psychopathol Behavioral Assess 25(2):105–119, https://doi.org/10.1023/A:1023394932263.
63. Proctor SP, Heaton KJ, Heeren T, White RF. 2006. Effects of sarin and cyclosarin exposure during the 1991 Gulf War on neurobehavioral functioning in US army veterans. Neurotoxicology 27(6):931–939, PMID: 16982099, https://doi.org/10.1016/j.neuro.2006.08.001.
64. Steele L, Sastre A, Gerkovich MM, Cook MR. 2012. Complex factors in the etiology of Gulf War Illness: wartime exposures and risk factors in veteran subgroups. Environ Health Perspect 120(1):112–118, PMID: 21930452, https://doi.org/10.1289/ehp.1003399.
65. Barth SK, Dursa EK, Bossarte RM, Schneiderman AI. 2017. Trends in brain cancer mortality among U.S. Gulf War veterans: 21 year follow-up. Cancer Epidemiol 50(Pt A):22–29, PMID: 28780478, https://doi.org/10.1016/j.canep.2017.07.012.
66. Heaton KJ, Palumbo CL, Proctor SP, Killiany RJ, Yurgelun-Todd DA, White RF. 2007. Quantitative magnetic resonance brain imaging in US army veterans of the 1991 Gulf War potentially exposed to sarin and cyclosarin. Neurotoxicology 28(4):761–769, PMID: 17485118, https://doi.org/10.1016/j.neuro.2007.03.006.
67. Chao LL, Rothlind JC, Cardenas VA, Meyerhoff DJ, Weiner MW. 2010. Effects of low-level exposure to sarin and cyclosarin during the 1991 Gulf War on brain function and brain structure in US veterans. Neurotoxicology 31(5):493–501, PMID: 20580739, https://doi.org/10.1016/j.neuro.2010.05.006.
68. Chao LL, Abadjian L, Hlavin J, Meyerhoff DJ, Weiner MW. 2011. Effects of low-level sarin and cyclosarin exposure and Gulf War Illness on brain structure and function: a study at 4T. Neurotoxicology 32(6):814–822, PMID: 21741405, https://doi.org/10.1016/j.neuro.2011.06.006.
69. Chao LL, Reeb R, Esparza IL, Abadjian LR. 2016. Associations between the self-reported frequency of hearing chemical alarms in theater and regional brain volume in Gulf War veterans. Neurotoxicology 53:246–256, PMID: 26920621, https://doi.org/10.1016/j.neuro.2016.02.009.
70. Chao LL, Zhang Y, Buckley S. 2015. Effects of low-level sarin and cyclosarin exposure on white matter integrity in Gulf War veterans. Neurotoxicology 48:239–248, PMID: 25929683, https://doi.org/10.1016/j.neuro.2015.04.005.
71. Yokoyama K, Araki S, Murata K, Nishikitani M, Okumura T, Ishimatsu S, et al. 1998. Chronic neurobehavioral and central and autonomic nervous system effects of Tokyo subway sarin poisoning. J Physiol Paris 92(3–4):317–323, PMID: 9789830, https://doi.org/10.1016/s0928-4257(98)80040-5.
72. Yanagisawa N, Morita H, Nakajima T. 2006. Sarin experiences in Japan: acute toxicity and long-term effects. J Neurol Sci 249(1):76–85, PMID: 16962140, https://doi.org/10.1016/j.jns.2006.06.007.
73. Yamasue H, Abe O, Kasai K, Suga M, Iwanami A, Yamada H, et al. 2007. Human brain structural change related to acute single exposure to sarin. Ann Neurol 61(1):37–46, PMID: 17187377, https://doi.org/10.1002/ana.21024.
74. U.S. General Accounting Office. 2003. Gulf War Illnesses: Preliminary Assessment of DOD Plume Modeling for U.S. Troops' Exposure to Chemical Agents (GAO-03-833T). https://www.gao.gov/assets/120/110012.pdf [accessed 19 February 2022].
75. Davies HG, Richter RJ, Keifer M, Broomfield CA, Sowalla J, Furlong CE. 1996. The effect of the human serum paraoxonase polymorphism is reversed with diazoxon, soman and sarin. Nat Genet 14(3):334–336, PMID: 8896566, https://doi.org/10.1038/ng1196-334.
76. Institute of Medicine. 2004. Gulf War and Health: Updated Literature Review of Sarin. Washington, DC: The National Academies Press. PMID: 25009884, https://doi.org/10.17226/11064 [accessed 11 April 2022].

77. Lash TL, VanderWeele TJ, Rothman KJ. 2021. Measurement and measurement error. In: *Modern Epidemiology*. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ, eds. Philadelphia: Wolters Kluwer, 291.
78. Costa LG, Giordano G, Cole TB, Marsillach J, Furlong CE. 2013. Paraoxonase 1 (PON1) as a genetic determinant of susceptibility to organophosphate toxicity. Toxicology 307:115–122, PMID: 22884923, https://doi.org/10.1016/j.tox.2012.07.011.
79. Dudbridge F, Fletcher O. 2014. Gene–environment dependence creates spurious gene–environment interaction. Am J Hum Genet 95(3):301–307, PMID: 25152454, https://doi.org/10.1016/j.ajhg.2014.07.014.
80. Rothman KJ. 1976. Causes. Am J Epidemiol 104(6):587–592, PMID: 998606, https://doi.org/10.1093/oxfordjournals.aje.a112335.
81. Rothman KJ. 2012. *Epidemiology: An Introduction*. 2nd ed. New York, NY: Oxford University Press.
82. Phillips PC. 2008. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9(11):855–867, PMID: 18852697, https://doi.org/10.1038/nrg2452.
83. Figueiredo TH, Apland JP, Braga MFM, Marini AM. 2018. Acute and long-term consequences of exposure to organophosphate nerve agents in humans. Epilepsia 59 (suppl 2):92–99, PMID: 30159887, https://doi.org/10.1111/epi.14500.
84. van Helden HP, Vanwersch RA, Kuijpers WC, Trap HC, Philippens IH, Benschop HP. 2004. Low levels of sarin affect the EEG in marmoset monkeys: a pilot study. J Appl Toxicol 24(6):475–483, PMID: 15558834, https://doi.org/10.1002/jat.1001.
85. Institute of Medicine 2003. Gulf War and Health: Volume 2. Insecticides and Solvents. Washington: National Academy Press. https://doi.org/10.17226/10628 [accessed 11 April 2022].
86. Office of the Special Assistant for Gulf War Illnesses. 1997. Information Paper: M8A1 Automatic Chemical Agent Alarm. https://gulflink.health.mil/m8a1alarms/ [accessed 19 February 2022].
87. Office of the Special Assistant for Gulf War Illnesses. 1996. Coalition Chemical Detections and Health of Coalition Troops in Detection Area. https://gulflink.health.mil/coalitn.html [accessed 19 February 2022].
88. Office of the Special Assistant for Gulf War Illnesses. 1998. Case Narrative: Czech and French Reports of Possible Chemical Agent Detections. https://gulflink.health.mil/czech_french/ [accessed 19 February 2022].
89. Bendtsen KM, Bengtsen E, Saber AT, Vogel U. 2021. A review of health effects associated with exposure to jet engine emissions in and around airports. Environ Health 20(1):10, PMID: 33549096, https://doi.org/10.1186/s12940-020-00690-y.
90. Haley RW. 1997. Is Gulf War syndrome due to stress? The evidence reexamined. Am J Epidemiol 146(9):695–703, PMID: 9366616, https://doi.org/10.1093/oxfordjournals.aje.a009343.
91. White RF, Steele L, O'Callaghan JP, Sullivan K, Binns JH, Golomb BA, et al. 2016. Recent research on Gulf War Illness and other health problems in veterans of the 1991 Gulf War: effects of toxicant exposures during deployment. Cortex 74:449–475, PMID: 26493934, https://doi.org/10.1016/j.cortex.2015.08.022.
92. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. 2014. Methodological challenges in Mendelian randomization. Epidemiology 25(3):427–435, PMID: 24681576, https://doi.org/10.1097/EDE.0000000000000081.